SCHMEISSER, MARIANNE G., M.S. Creation of a GPR18 Homology Model Using Conformational Memories. (2013)
Directed by Dr. Patricia H. Reggio. 46 pp.

G-protein coupled receptors (GPCRs) make up the largest family of eukaryotic membrane receptors, covering a broad range of cellular responses in the body. This wide range of activity makes them important pharmacological targets. In general, all Class A GPCRs share a common structure that consists of seven transmembrane alpha helices, connected by extracellular and intracellular loops, an extracellular N-terminus, and an intracellular C-terminus. These similarities can be used to construct a model of an unknown receptor, which can then be used to help guide further studies of this receptor and its pharmacology.

The orphan GPCR GPR18 is a member of the Class A subfamily of GPCRs. GPR18 binds both lipid-like and small molecule ligands, such as NAGly and abnormal-cannabidiol (Abn-CBD), leading to belief that GPR18 may be the Abnormal Cannabinoid Receptor. The goal of this project was to construct a model of GPR18 in its inactive state and to explore the binding site of a key antagonist already identified for this receptor. A model of the GPR18 inactive (R) state was created using the μ-Opioid receptor (MOR) crystal structure as template (PDB: 4DKL). The Monte Carlo/simulated annealing method, Conformational Memories (CM), was used to study the accessible conformations of three GPR18 transmembrane helices (TMHs) with important sequence divergences from the MOR template: TMH3, TMH4, and TMH7. CM was also used to calculate the accessible conformations for TMH6, which allowed the choice of TMH6 conformers appropriate for the GPR18 R and R* models.

Docking studies were guided by the hypothesis that a positively charged residue (either R2.60 or R5.42) may be the primary ligand interaction site in the GPR18 binding pocket. The binding pocket of the antagonist, cannabidiol (CBD) was explored in the inactive state GPR18 model using Glide, an automatic docking program in the Schrödinger modeling suite. These studies suggested that both of these arginines are primary interaction sites for CBD. With the pocket determined, extracellular and intracellular loops were calculated using another Monte Carlo technique, Modeler. Once loops were attached, the N and C termini were modeled and added as well. The N terminus displayed a small helical portion that lay atop the bundle. Together with the EC2 loop, this N terminus closes off the EC domains of GPR18 from the extracellular milieu. This result is consistent with the structure of the S1PR1 receptor, which is closed to the extracellular milieu, but allows ligands to gain entry via the lipid bilayer. With the identification of key residues and a complete GPR18 bundle, further mutation studies and dynamic simulations can be used to further refine and test these modeling results.

CREATION OF A GPR18 HOMOLOGY MODEL USING COMFORMATIONAL MEMORIES

by

Marianne G. Schmeisser

A Thesis Submitted to
the Faculty of the Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Greensboro
2013

Approved by
Patricia H. Reggio_____
Committee Chair

APPROVAL PAGE

This thesis written by MARIANNE G. SCHMEISSER has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Chair Patricia H. Reggio_____

Committee Members Mitchell Croatt_____

Ethan Taylor_____

July 26, 2013_____
Date of Acceptance by Committee

July 26, 2013_____
Date of Final Oral Examination

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER I

INTRODUCTION

Background:

Microglia are the first, and main, form of active immune defense in the central nervous system (CNS), and possess the ability to change receptor expression to execute rapid, pointed migration towards affected tissue. Deregulation of this migration cycle and cellular change leads to proinflammatory and cytotoxic responses, recognized in several neurodegenerative diseases, such as multiple sclerosis and Alzheimer's.[1] In other parts of the body, like the endometrium, this cellular migration is believed to play a large part in the painful, and sometimes infertility causing, condition of endometriosis.[2] The endocannabinoid system has been shown to regulate microglial migration[1], and recent cell migration studies have found a possible new target protein, GPR18, at which the endocannabinoids may exert their immunomodulatory effects.[2,3,4] GPR18 has recently been found in primary melanoma cells and in a human endometrial cell line, HEC-1B.[2] This pattern of expression suggests GPR18 as a possible therapeutic target for inflammatory diseases, such as endometriosis, as well as a target for neurodegenerative diseases such as multiple sclerosis and Alzheimer's disease.

GPR18 belongs to the Class A, or Rhodopsin-like, subfamily of G-Protein Coupled Receptors (GPCRs). GPCRs are transmembrane proteins that represent the

largest family of eukaryotic plasma membrane receptors, responsible for the majority of signal transduction across cell membranes. GPCRs respond to many factors, such as hormones, neurotransmitters, odorants, and light, making them important pharmacological targets.[5] There are four distinguishing characteristics of all Class-A GPCRs, shown in Figure 1:

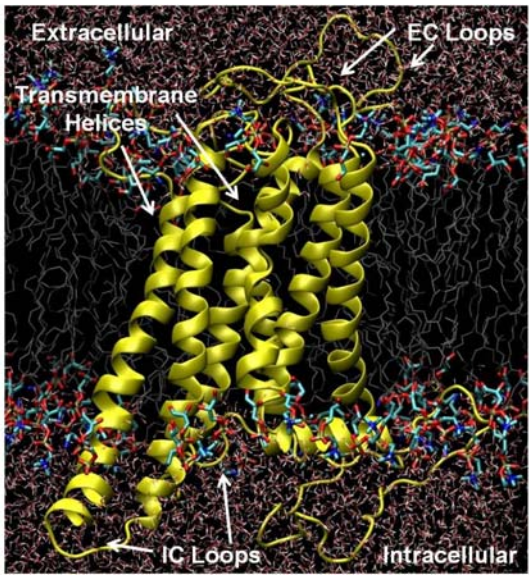


Figure 1. Position of Typical GPCR Within a Cell Membrane

1. Seven transmembrane α-helices (TMHs)
2. An extracellular N-terminus
3. Extracellular (EC) and Intracellular (IC) loops that connect each of the seven TMHs

4. An intracellular C-terminus that begins with a helical section, termed Helix 8, that lies parallel to the cell membrane (H8)

With all GPCRs containing these same structural features, this makes comparing the differences all the more important. However, Class A receptors vary in the total number of amino acids in their sequences. To make comparisons easier, a new numbering system was constructed to describe the transmembrane sections, as the main length variances appear in the connecting loops. The numbering system used here is that described by Ballesteros and Weinstein in 1995.[6] Here, the most highly conserved residue in each TMH is assigned the designation of x.50. The x is replaced by the helix number and the descriptor can be followed by the sequence number in parentheses. All other residues within the TMH are numbered relative to this x.50 residue. As an example, the most conserved residue in TMH2 is aspartic acid D2.50. The residue immediately before is V2.49 and immediately after is L2.51 in the GPR18 sequence.

For many years, the rhodopsin crystal structure (PDB: 1GZM)[7] was used as the template for the creation of GPCR homology models. Recently, x-ray crystal structures of other Class A (Rhodopsin-like) GPCRs have become increasingly available. These include meta-rhodopsin II[8], the β2-adrenergic receptor (β2-AR)[9], the β1 adrenergic receptor (β1-AR)[10], the adenosine A2A receptor[11], the CXCR4 receptor[12], the dopamine D3 receptor[13], the histamine H1 receptor[14], S1PR1 receptor[15], the nociception/orphanin FQ receptor[16] and the μ, Δ and κ opioid receptors.[17,18,19]

Of all of the recently available GPCRs crystal structures, the μ-opioid receptor exhibits the highest sequence homology (54%) with GPR18 in key transmembrane helix segments. Unlike the majority of GPCRs, for which residues within the binding pocket and the second EC loop partially obscure the ligand, μ has a largely exposed extracellular opening, shown in Figure 2 in comparison with Rhodopsin.[17]

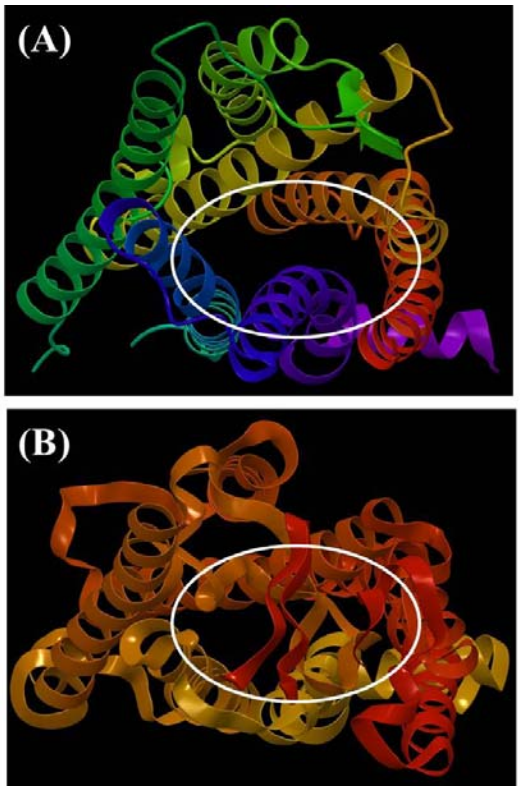


Figure 2. Comparison of Extracellular Area between
(A) μ-Opioid and (B) Rhodopsin

Because GPR18 has been shown to recognize neutral, lipophilic ligands and small molecules typically bound by the cannabinoid receptors,[2, 20] it is likely that GPR18 will exhibit a closed extracellular domain with access to the binding pocket from within the lipid bilayer. This type of entry has been proposed both for the S1PR1 receptor[15] and the cannabinoid CB_2 receptor.[21,22]

GPCR Activation

Information about GPCR activation has come mainly from biophysical studies on rhodopsin[23] and the beta-2-adrenergic receptor.[24,25] Within each Class A GPCR binding pocket, there is thought to be a set of residues that change conformation upon agonist binding. These are called "toggle switch" residues and typically include residue W6.48 of the TMH6 CWXP motif and another residue that interacts with W6.48. Thus, when an antagonist binds, the ligand either directly blocks 6.48 from changing conformation, or causes other steric hindrances that indirectly halt a conformational change. In the inactive state of rhodopsin, the $\chi 1$ dihedral of W6.48 has been shown to be -60°, g^+.[7]

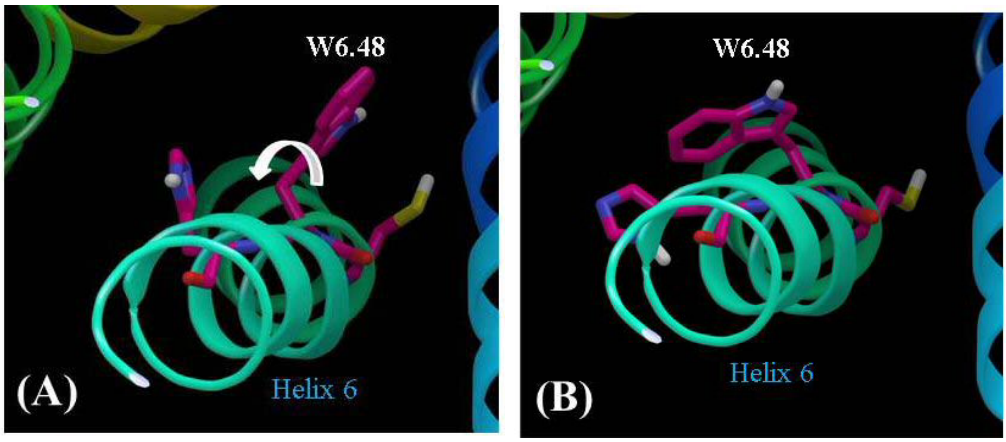


Figure 3. Toggle Switch between (A) Inactive and (B) Active States

This conformation is due to the close proximity of the beta-ionone ring of the covalently bound ligand, 11-cis-retinal, to W6.48, which locks the tryptophan in this conformation. When light activates rhodopsin, the ligand isomerizes into all-trans retinal. This conformational shift causes the beta-ionone ring to move away from TMH6 and towards TMH4, allowing W6.48 to flip into a *trans* χ1 conformation, 180°. This change allows TMH6 to flex in the CWXP hinge region and straighten.[25] This straightening, in turn, breaks the "ionic lock" between R3.50 and E/D6.30 at the intracellular end of the receptor. The result is the formation of an intracellular opening of the receptor, exposing residues that can interact with the C-terminus of the Gα sub-unit of the G protein.[26] The μ-opioid "ionic lock" occurs in a slightly different position, between R3.50 and T6.34[27], so it is possible that GPR18 exhibits an ionic lock with a different partner to R3.50, such as S6.33.

Ligands of GPR18

GPR18 is similar to the cannabinoid receptors in that it binds to both small molecule and lipid-like ligands, shown below and in Appendix 1. In 2006, Kohno et al identified eicosa-5,8,11,14-tetraenoylamino-acetic acid (N-arachidonoylglycine; NAGly) as the endogenous GPR18 ligand. NAGly is structurally similar to the CB_1 endogenous ligand N-arachidonoyl ethanolamine (AEA).[20] The difference between the two lies in the head group region, where AEA contains an ethanolamine versus NAGly's glycine. This difference renders NAGly inactive at both the CB_1 and CB_2 receptors. While the biological functions of NAGly are not very well understood, it has been shown to have some analgesic properties similar to AEA, via inhibition of the hydrolytic activity of fatty acid amid hydrolase (FAAH) on AEA. However, due to the level of expression of NAGly in a variety of tissues (e.g. skin, small intestine, kidney, testis, and brain) it likely is involved in other physiological functions. Arachadonic acid-derivatives usually play a major role in inflammation and pain, as seen in the cannabinoid receptors.[20]

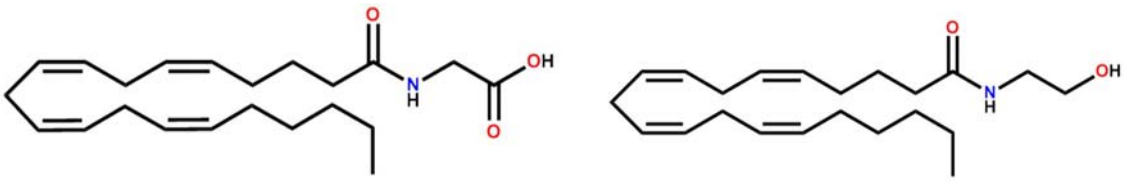


Figure 4. Endogenous GPR18 ligand NAGly (left) compared to AEA (right)

GPR18 can also be activated by the atypical cannabinoid Abn-CBD, as well as several other cannabinoids, leading to the idea that GPR18 could be considered another cannabinoid receptor.[1]

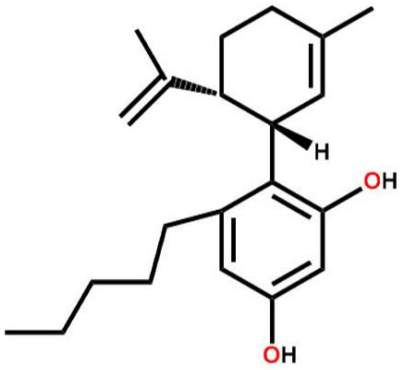


Figure 5. Structure of Abn-CBD

Cell migration assays have shown that GPR18 couples to $G_{i/o}$ proteins, with an agonist profile including Abn-CBD, O-1602, Δ^9-THC, and NAGly. The observed migration can be blocked by O-1918, arachidonyl serine, or cannabidiol (CBD). This ligand profile is similar to that documented for a previously postulated cannabinoid receptor, the Abn-CBD receptor, which was never cloned. There is increasing evidence, however, that GPR18 may in fact be the Abn-CBD receptor.[1-2]

CHAPTER II

HYPOTHESIS & METHODS

Goals:

This project focused on the computational creation of a homology model of the GPR18 R (Inactive) state using the TMH region of the recent μ-opioid receptor x-ray crystal structure[17] as a template. Because GPR18 has important sequence divergences from the template in TMH3, 6, and 7, Conformational Memories (CM), the Monte-Carlo/simulated annealing technique, was used to calculate new conformations for these helices. A possible binding pocket was identified, and the small molecule cannabidiol (CBD) was docked in the resultant model. Interactions surrounding the binding of this antagonist were analyzed to determine key binding pocket residues. The main goal was to create a usable computational model to help guide further investigation into this new receptor through mutation and molecular dynamics studies.

Methods:

Rotameric States in Amino Acid Side Chains

The rotamer nomenclature used herein is the same as that described by Shi et al in 2000[24], and the basic structure shown in Figure 6.

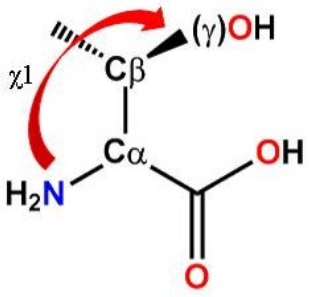


Figure 6. The χ1 angle displayed using Threonine

A *trans* χ1 angle is defined as when the heavy atom at the γ position lies opposite the backbone nitrogen (180°), when viewed from β-carbon to α-carbon. *Gauche*$^+$ (g^+) is defined as when the γ heavy atom is opposite the backbone carbon and has an angle of -60°, while *gauche*$^-$ (g^-) has the γ-carbon heavy atom opposite the α-hydrogen at an angle of +60°, viewed along the same axis.

Sequence Alignment

As discussed above, the orphan receptor GPR18 belongs to the Class-A, or Rhodopsin-like, subfamily of GPCRs. GPR18 contains the majority of the conserved residues of Class A GPCRs in TMHs 1, 2, 4, and 5 (N1.50, D2.50, W4.50, and P5.50), as well as the DRY motif in TMH3. However, it also contains some notable differences, including: 1) a substitution at position 6.48 (CFXP instead of CWXP), and 2) the nonconservative motif DVILY instead of the TMH7 NPXXY motif. In addition, TMH3 in GPR18 contains a helix-altering proline residue at position 3.36. The initial GPR18

model was constructed by first aligning the sequence with that of other Class A GPCRs using the highly conserved sequence motifs or residues as the alignment guides, shown in Appendix 2. The base crystal structure for the μ-opioid receptor (PDB: 4DKL) was mutated to the corresponding GPR18 sequence using Maestro (a Schrödinger Inc. program). Helices that contain helix deforming residues such as prolines or glycines were then studied using Conformation Memories (see below) to determine possible low free energy conformations, and then an appropriate substitute helix was chosen to incorporate into the model.

Conformational Memories (CM)

In order to explore the different conformations of TMH3, TMH6, and TMH7 dictated by their sequence divergences from the template structure, the conformational memories (CM) method was used. This method employs multiple Monte Carlo/simulated annealing random walks and the CHARMM (**C**hemistry at **HAR**vard **M**olecular **M**echanics) force field.[28] The CM method is widely used because it converges in a practical number of steps and it is capable of overcoming energy barriers efficiently. The CM method can fully characterize the conformational properties of a helix based on their free energies. This includes both the intrinsic energy of each conformational state and the probability that the helix will adopt a specific conformation relative to the other accessible conformations in an equilibrated thermodynamic ensemble. The calculations are performed in two phases, the Exploratory and Biased Annealing Phases.

Exploratory Phase

In the CM method, a random walk is used to identify the region of conformational space most probable for each torsion angle and bond angle. The initial temperature for each run is 3000 K with 50,000 Monte Carlo steps applied to each torsion or bond angle variation with cooling in 18 steps to a final temperature of 310 K. Each step consists of varying two dihedrals angles and one bond angle chosen at random from the entire set of variable angles. Each move is the accepted or rejected using the Metropolis criterion.[29] Accepted conformations in the Exploratory Phase are used to create "memories" of torsion angles and bond angles that were accepted. This information provides a map of the accessible conformational space of each TMH as a function of temperature.

Biased Annealing Phase

In the second phase of the CM calculations, the only torsion angles and bond angles moves attempted are those that would keep the angle in the "populated conformational space" mapped in the exploratory phase. The biased annealing phase begins at 749.4 K cooling to 310 K in 7 steps. Finally, 120 structures are output at 310 K for each TMH studied.

Identification of TMHs for CM Study

For the GPR18 model, the base μ-Opioid model required CM studies of three of helices: TMH3, TMH6, and TMH7. TMH3 would differ due to a proline at position 3.36,

which is unmatched in the μ sequence. For TMH6, GPR18 has a CFXP motif in TMH6, rather than the typical Class A CWXP motif as in the μ-opioid receptor. The difference of the W→F change at 6.48 and the added bulk of a T→M change at 6.49 would induce a change in the helix conformation, which would be amplified further by the need to accommodate the following bulk of an F at 6.51 and an H at 6.52 in GPR18. Lastly, calculation of a new conformation for TMH7 was necessitated by the lack of the highly conserved P7.50 in GPR18, as well as an altered motif, DVILY rather than the more conserved NPXXY in the μ-opioid receptor.

CM Study of GPR18 TMH3, TMH6, and TMH7

Models of TMH3, TMH6, and TMH7 were created using the CM technique. The backbone dihedrals of each α-helix were set to the standard φ (-63°) and ψ (-41.6°) for TMHs as described by Ballesteros and Weinstein in 1995.[6] Our established protocol is to allow all torsion angles to vary ±10°, and to allow a larger variation of ±50° in regions containing helix bending residues such as prolines, glycines, serines and/or threonines.[30] Individual bond angles were allowed to vary ±8°, and the following were allowed to vary ±15°: C-O-H on Ser, Thr and Tyr; C-N-H on Trp, His, Gln, Asn and Lys; C-S-H on Cys and C-S-C on Methionine.[28] These bond angle variations are used since certain residues and sulfur atoms can accommodate bond angle changes, while aliphatic hydrogen bond angles were not varied.

TMH3: All GPCRs were aligned using the conserved motif of E/DRY in TMH3. GPR18 has a unique proline at position 3.36. No other Class A GPCR has a proline at this position, so the bend introduced into the helical structure has no previous structural match. To allow for helix distortion, the region of i (P3.36) to i-4 (T3.32) was considered "flexible," with the phi and psi values in this area allowed to vary by ±50°.

TMH6: TMH6 contains a very similar flexible motif, CFXP, compared to the Class A GPCR conserved motif of CWXP. However, GPR18 has a methionine at position 6.49, which is a fairly bulky residue. There are also a couple of bulky residues after the CFMP region, which would have to be accommodated by the new output. Previous studies show that W6.48 is a "toggle switch" for activation, χ^1 is in a g+ position in the R state, and in trans in the R* state.[24] Conformational Memories outputs were used to select an appropriate helix for an R and R* model. In TMH6 CM calculations, the i (P6.50) to i-4 (V6.46) region was allowed to vary by ±50°.

TMH7: The μ-opioid receptor contains the Class A GPCR conserved motif of NPXXY. However, GPR18 contains neither the conserved P7.50 nor the motif of NPXXY. To account for possible changes in conformation due to this sequence variance, the flexible region of i (V7.50) to i-4 (T7.46) was chosen and allowed to vary by ±50° in the Conformational Memories run.

Superimposition and Determination of Helices

The base template bundle of the μ-opioid receptor had its sequence mutated to that of GPR18. The resultant initial model was then was pulled apart by 2Å to allow room for the side chains to accommodate each other in the new bundle. For each Conformation Memories study, all 120 output conformations were superimposed onto the corresponding template helix in the initial GPR18 model. For all superimpositions described, the Cα's of the sequence up to the flexible region were used as the basis for the superimposition.

Once the outputs were superimposed upon the initial model, an appropriate helix was chosen after eliminating helices that had steric clashes with the binding pocket. Helix 3 was superimposed from the intracellular residues up to S3.37, the residue immediately preceding the proline. With the conformational change in TMH6 between R and R* states, the CM outputs were sorted based on the χ1 angle of F6.48. The g^+ (-60° range) were superimposed onto the base helix at the intracellular end, from K6.30 to L6.45; while those with the χ1 angle *trans* (for the R* model), the outputs were superimposed on the extracellular end of the temple helix, from F6.51 to G6.61. For TMH6, the R helix chosen was based on the presence of F6.48 χ1 in a g+ position, and ability to form the TMH3/TMH6 ionic lock (R3.50/S6.33 for GPR18) characteristic of the inactive state. The chosen R* helix fit the requirements that the TMH3/TMH6 ionic lock must be broken and that the χ1 angle of F6.48 be in a *trans* position. For TMH7, the

outputs were superimposed on the extracellular region down to S7.45, the residue immediately before the i-4 residue of T7.46.

Ligand Docking Protocol

Before docking cannabidiol (CBD), an AM-1-conformational search was performed to identify its global minimum energy conformation. The global min was then used for docking studies via the automatic docking program, Glide (Schrodinger Inc.). Glide generates a grid based on the centroid of select residues in the binding site (from the manual dock).[31] In GPR18, R5.42 was used as the primary interaction site to help position CBD in an antagonist position over 6.48. The resulting ligand/receptor complex was energy minimized using CHARMM, in the same process as described below.

Minimization Protocol

Each bundle was pulled apart 2Å away from a central point to allow room for the side chains to accommodate each other in the new bundle. Once the new helices were chosen and inserted into the receptor bundle, each residue was adjusted manually to its most energetically favorable position, while allowing room for all other residues and preventing any Van der Waals conflicts. The $\chi 1$ torsion angles of C6.47, W6.48 and H6.52 were adjusted to *trans*/g^+/g^+ respectively to be in agreement with the "toggle switch" proposed by Shi and co-workers.[24] Also, R3.50 and S6.33 were oriented towards each other to promote the formation of the ionic lock when the bundle is minimized. The

distance between the first nitrogen of the guanidine group of R3.50 and the hydroxyl oxygen of S6.33 was constrained to a distance of 2.8Å ± 0.4Å with a force of 15kJ/mol. Once all Van der Waals overlaps had been relieved, the minimization was started using an OPLS_2005 force field, no solvent, and a distance dependent dielectric. All backbone phi / psi dihedrals were constrained with a force of 500kJ/mol, while the amino acid side chains of each residue were allowed to vary. Minimizations were performed with the docked ligand CBD present, as previous attempts of minimizing the bundle in the absence of ligand resulted in collapse of the binding pocket.

Loop and Termini Methodology

Once appropriate substitute helices for TMH3, 4, 6, and 7 were chosen, loop segments were built and added to the models using Modeler. This Monte Carlo technique developed by Fiser and co-workers uses a template library of possible side chain conformations from the Protein Data Bank for all amino acids.[32] Using the CHARMM force field, each loop was varied and assigned an objective function ranking value.[28] This value is based on steric interactions and hydrogen bonding of each possible conformation, and the 250 lowest energy loop outputs were used for further analysis and selection.

Similar to the extracellular and intracellular loops, the N and C termini were built and added to the GPR18 model once the loops were complete. The same Monte Carlo

technique was used to calculate possible side chain conformations. The 250 lowest energy loop outputs were then sorted and analyzed to determine a final structure.

Interaction Energy Calculation

The interaction energy between the ligand, CBD, and the minimized receptor complex was calculated using a Molecular Mechanics tool in Maestro (Schrodinger, 2006). The OPLS_2005 force field was used, with no solvent and a distance dependent dielectric. The ligand was identified manually (selected by hand), and all residue interactions within 5Å of the ligand were used in the calculation.

CHAPTER III

RESULTS

GPR18 Model based on the μ-Opioid Receptor:

An inactive (R) receptor bundle was originally constructed using the μ-opioid receptor structure as a template. Upon studying the CM output structures for TMH3, it was observed that both TMH3 and 4 in the μ-opioid receptor were significantly bent, highlighted by the red dashed lines in Figure 7.

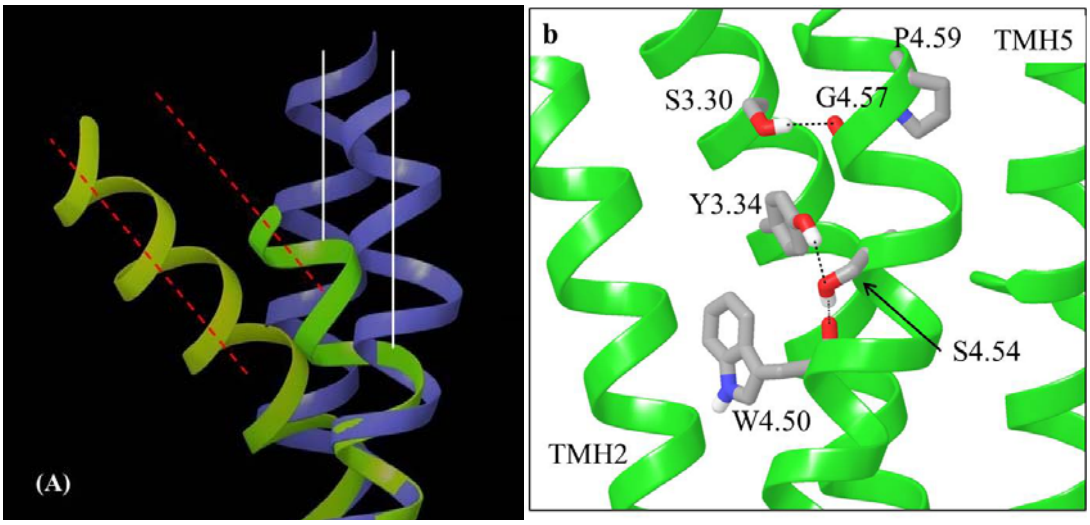


Figure 7. (A) Comparison of the EC ends of TMH3 and 4 between MOR (yellow/green) and Straight Helices (blue). (B) Close-up of residues forming the hydrogen bonding network in MOR

This bend is due to a hydrogen bonding network between TMH3 and 4, which is unmatched in the GPR18 sequence. Using the μ-opioid receptor TMH3 as a base for the

GPR18 TMH3 would introduce a bend bias into the structure that would not fit the kink induced by the proline at 3.36.

TMH6

Despite GPR18's CFMP motif versus the conserved CWXP motif of other Class-A GPCRs, the rest of the helical sequence of TMH6 does not differ substantially from that of the μ-opioid receptor template. Thus, the CM output chosen for the inactive, or R, bundle has a similar conformation to that of the μ–opioid receptor TMH6, with a wider bend to accommodate for the bulk of GPR18's CMFP region as well as the bulk of F6.51 and H6.52. The similar EC position of the helices also matches the similar loop lengths of both GPR18 and the μ-opioid receptor crystal. However, upon docking and minimization, it was observed that the chosen helix could not pack well with TMH5 and TMH7. Therefore, a second TMH6 output was chosen, which allowed for better packing with TMH7.

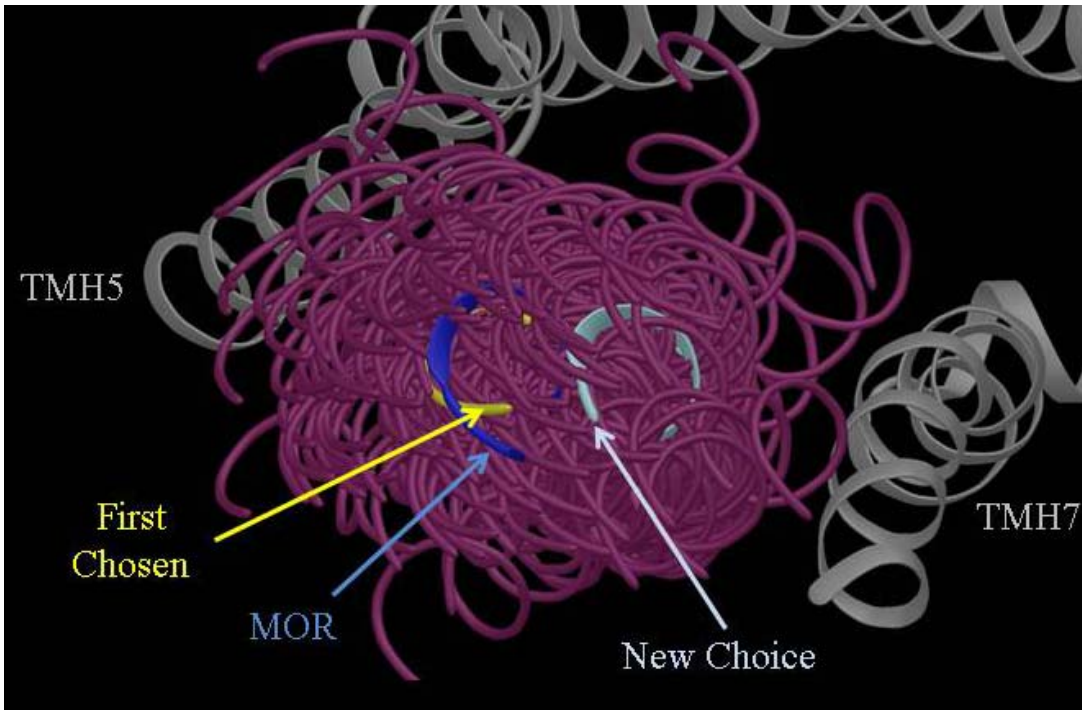


Figure 8. Extracellular View of TMH6 and labeled choices

As the TMH6 for the inactive bundle was superimposed on the intracellular side, S6.33 remained in a position to promote a possible ionic lock between it and the conserved residue R3.50, shown in Figure 9.

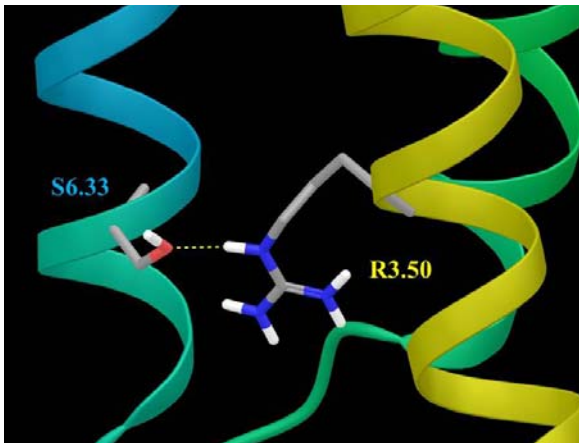


Figure 9. Ionic lock between R3.50 and S6.33

GPR18 Model – Ideal Helix Base Structures:

The μ-opioid receptor crystal structure template has two major differences with GPR18 concerning TMH3. First, GPR18 contains a proline at the position 3.36 that is unmatched in the μ-opioid receptor sequence, and second, the μ-opioid receptor bundle is built to accommodate bulky ligands that enter extracellularly. Despite varying the flexible region around the proline in the GPR18 model, the angles did not compensate for the bend already built into the template μ-opioid receptor TMH3. For this reason, TMH3 was built as an ideal helix and input to Conformational Memories using the same flexible region, from T3.32 to P3.36.

TMH3

Figure 10 shows the TMH3 output structures (blue) superimposed intracellularly onto the mutated μ-opioid receptor base structure. The red helix illustrated in Figure 10 was the helix chosen for the GPR18 bundle.

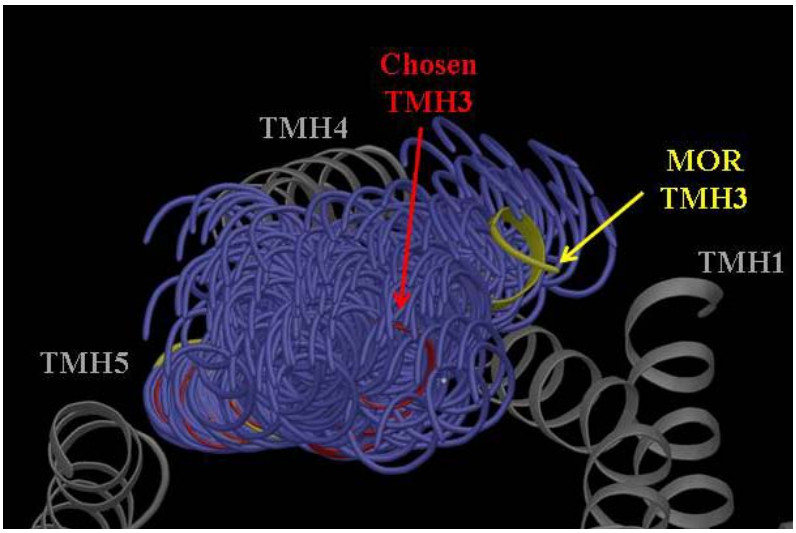


Figure 10. Extracellular view of TMH3 outputs and the chosen helix (red)

As mentioned before, the MOR crystal has a very open extracellular end to permit the docking of large ligands such as morphine and naloxone, which have structures very different from that of CBD and NAGly. Thus, from the span of outputs, a helix was required that was straighter, to help constrict the extracellular side of the bundle. However, the helices further into the binding pocket were not chosen as we did not want to completely constrict the binding pocket once the bundle was minimized. Also, there are aromatic residues on the top of both TMH2 and 5, which restrict how far to either side

the helix can be positioned. Lastly, due to the proline at 3.36, a TMH3 with a face shift needed to be chosen so that C3.25 at the EC end of TMH3 was positioned to form a disulfide bridge with the EC2 loop Cys(172). This disulfide bridge is found in nearly all Class A GPCRs.

TMH4

TMH4 in the μ-opioid receptor crystal structure is bent due to a hydrogen bonding network it shares with TMH3 in the μ structure, shown previously in Figure 7. This hydrogen bonding network pulls TMH4 into a bent position to match the significant bend in TMH3. As GPR18's TMH3 was already altered due to the unique P3.36, TMH4 in GPR18 lacks any reason to be bent in the same way as in the μ-opioid receptor. So a CM study of TMH4 was undertaken using an ideal helix as the starting structure. Comparing the GPR18 sequence with the μ-opioid receptor template resulted in the addition of another turn onto TMH4, so that the helix transitioned to loop at K4.65. Using an ideal helix as the starting structure for TMH4 permitted a choice for TMH4 that worked with the chosen GPR18 TMH3.

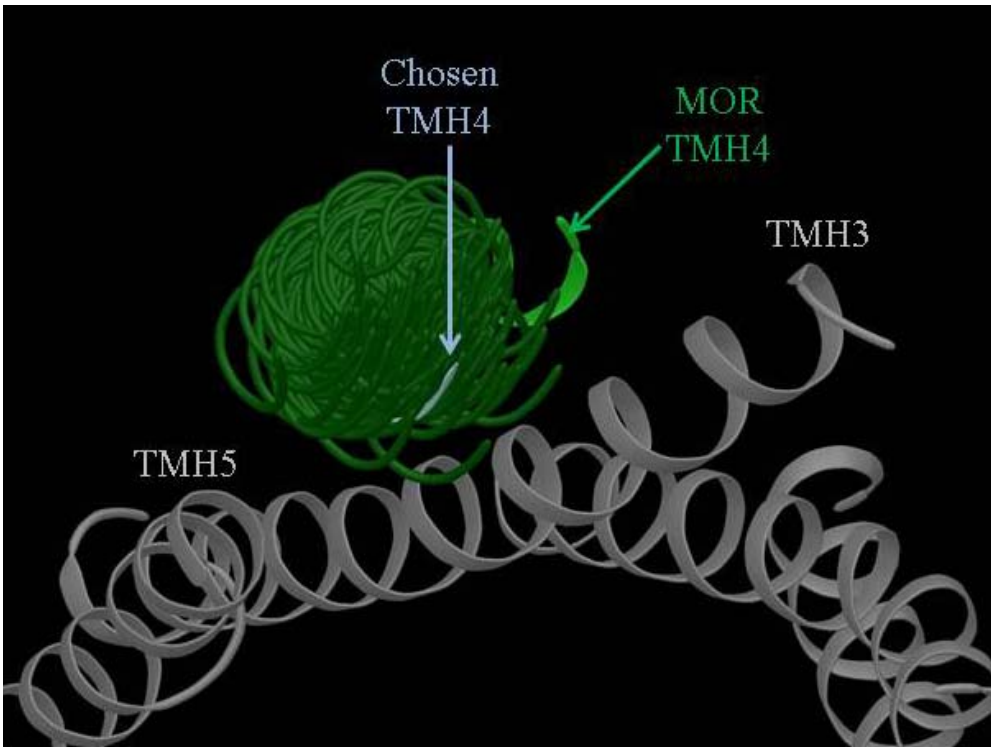


Figure 11. Extracellular view of TMH4 outputs and the chosen helix (light blue)

TMH7

Unlike the majority of Class A GPCRs, GPR18 does not contain the conserved motif of NPXXY within TMH7, so the calculated output would likely have a much different bend in it due to the lack of a proline. From the spread of outputs shown in Figure 12, there are no outputs that are similar to that of the MOR crystal base (as would be expected).

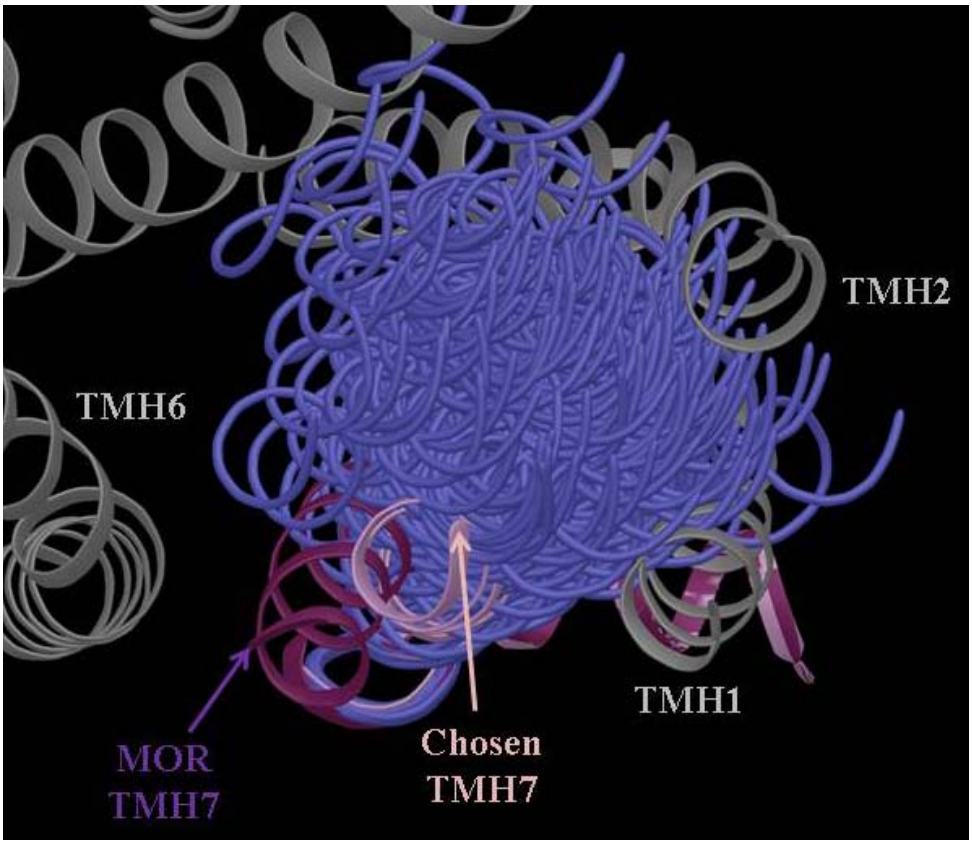


Figure 12. Extracellular view of TMH7, with the chosen helix (pale purple)

When placed into the bundle, steric clashes with aromatic residues on the interface between TMH1 and 7 and TMH6 and 7 could be used to eliminate helices. The helix chosen, shown in a light purple in Figure 12, had no steric clashes with the rest of the bundle, and packed well with the new TMH6 that had been chosen earlier.

Ligand Conformational Searches:

O-1918

From studies by McHugh,[1] the small molecule O-1918 was seen to antagonize the cell migration induced by NAGly. Using this information, O-1918 was chosen as the antagonist to be used for docking studies in the newly developed GPR18 R model. Using the program Spartan, the lowest energy conformation (or global min) was calculated using an AM-1 conformational search.

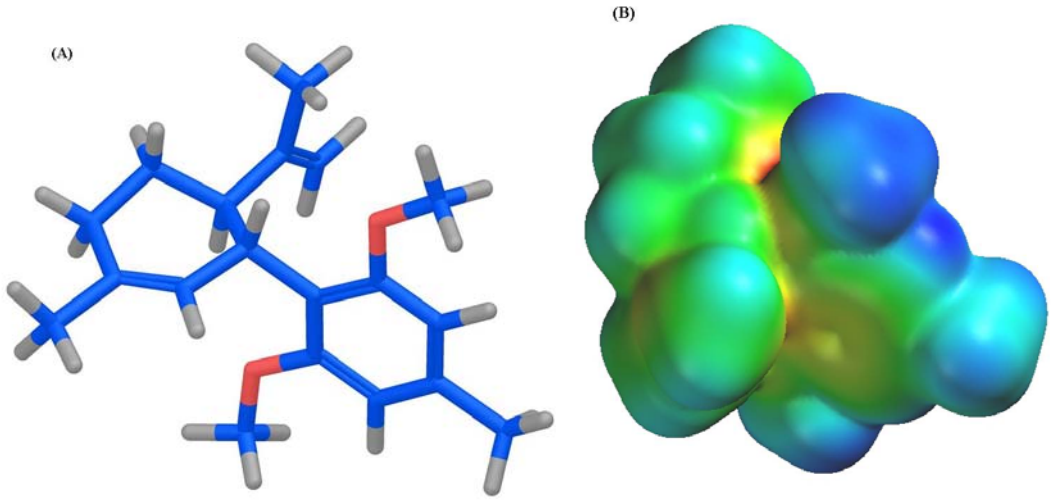


Figure 13. (A) O-1918 Global Minimum Conformation. (B) Electrostatic Potential Surface of O-1918 Global Min.

From the electrostatic potential map shown in Figure 13, the methoxy oxygens appeared to be heavily occluded by the rest of the molecule, so we hypothesized that a cation-pi interaction between the aromatic ring of O-1918 and one of the two arginines facing the binding pocket would be the primary interaction sites. An interaction with R2.60 left F6.48 with the ability to freely rotate, so R5.42 was chosen as the primary interaction

site. The Glide program, Maestro (Schrodinger, 2006), calculated few conformations that could work as a cation-pi, but there were problems minimizing the bundle and maintaining the docked position. Later information (unpublished) led to the conclusion that O-1918 could actually be working as an agonist within GPR18,[33] and we therefore abandoned the O-1918 study to look for another antagonist was chosen to be docked.

Cannbidiol (CBD)

The small molecule, Cannabidiol (CBD), was chosen as the new antagonist to be modeled and docked within the inactive GPR18 model.[2] Using the same procedure described above in the Methods section for O-1918, the global minimum conformation of CBD was calculated. The results are summarized below in Figure 14.

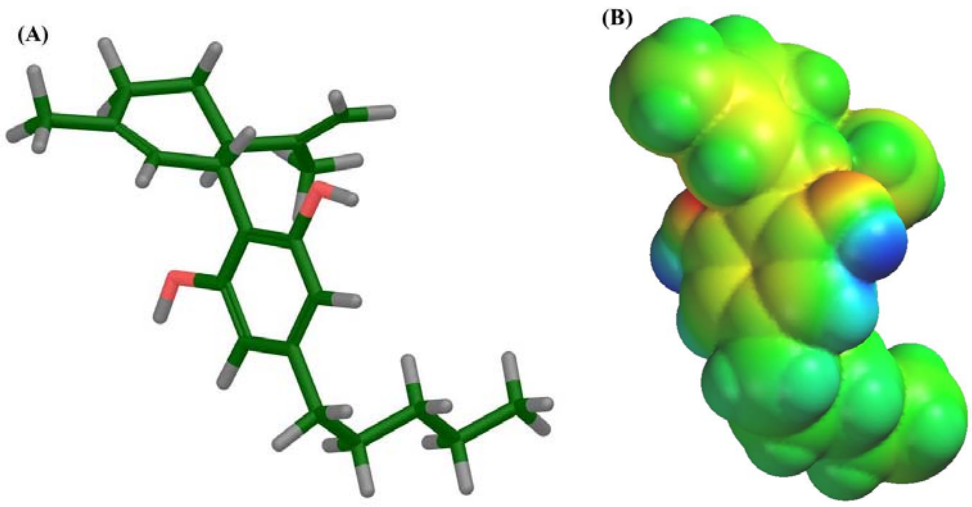


Figure 14. (A) CBD Global Minimum Conformation. (B) Electrostatic Map of CBD Global Min.

As can be seen in the electrostatic potential surface, shown in Figure 14B, the oxygens of the hydroxyl groups are more accessible than those of O-1918's methoxys. Thus, instead of cation-pi interactions, hydrogen bond interactions were more likely. When initially input into Glide, R5.42 was hypothesized to be the primary interaction site because the bulk of the ligand would impede the rotation of F6.48. An orientation of the ligand was observed that could possibly also reach R2.60, once the bundle was minimized around CBD.

As described in the Minimization Protocol (see the Methods section), the GPR18 bundle was minimized with CBD placed inside to maintain the binding pocket and prevent the bundle from collapsing on itself. The ligand was held rigid, aside from small rotations of the hydroxyl groups to help maintain the hydrogen bond with R5.42 and possibly to find a hydrogen bond with R2.60. Figure 15 shows the final docked position of CBD from an extracellular view, with the two arginines bracketing the ligand.

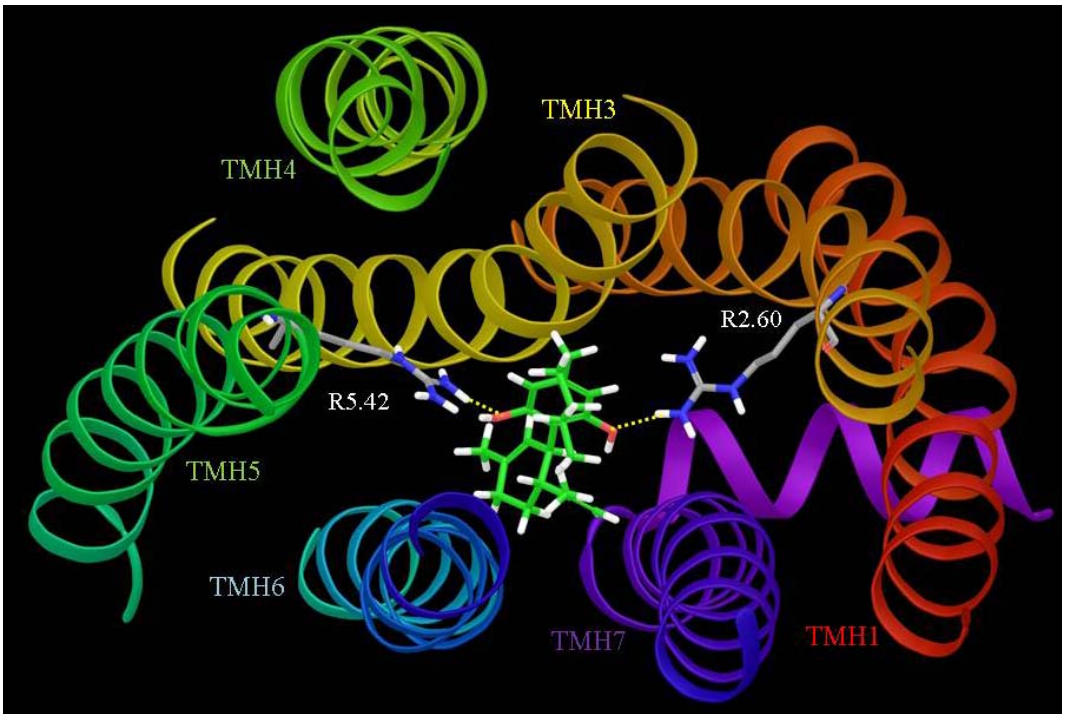


Figure 15. EC View of CBD docked position

When seen from the side, through TMH1 and 7, CBD (in green) sits close to 6.48 (shown in grey) so that when seen in Van der Waals, there is not enough room for the phenyl side chain of F6.48 to rotate into an active conformation.

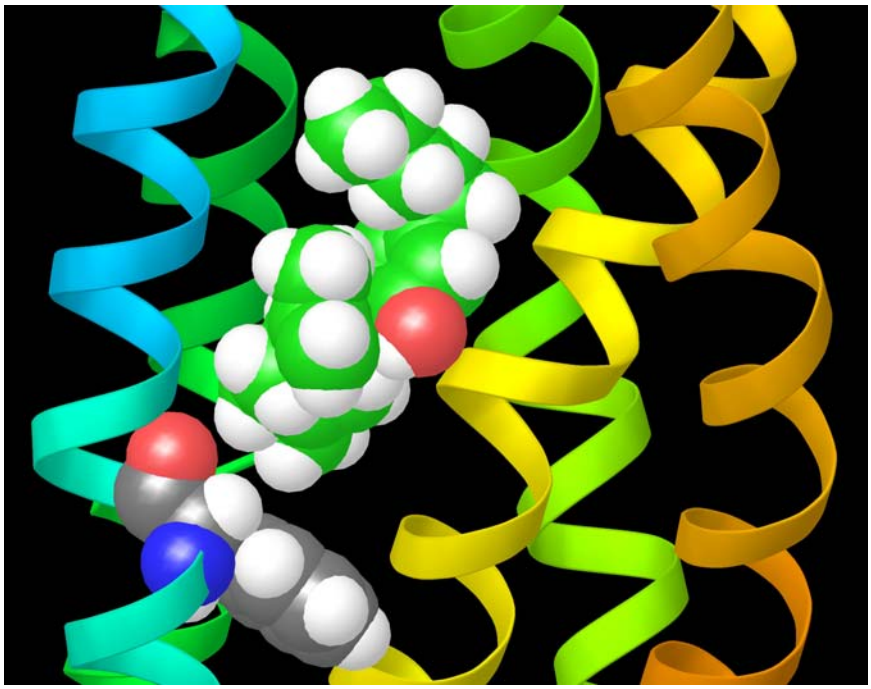


Figure 16. Docked CBD (green) sitting over F6.48 (grey)

Loop Construction and Modeling:

Extracellular (EC) Loops

Once the loops were built in Maestro, they were connected to the bundle in one entry. Modeler was used to calculate possible conformations for all three extracellular loops, in a high distance dependent dielectric (around 80) to simulate the aqueous environment surrounding the loops. GPR18 has small loops, with the EC1 and EC3 having only five and six residues respectively. The longer EC2 loop contains a disulfide bridge between C(172) and C3.25, which was the only structural constraint specified. One characteristic seen in multiple crystal structures is that the residue two residues after the disulfide Cys points down into the bundle. This was used as a screening tool to select

an appropriate EC2 loop conformation. Despite no secondary structure being specified, Modeler calculated a beta-sheet within the long EC2 loop, as can be seen in Figure 17.

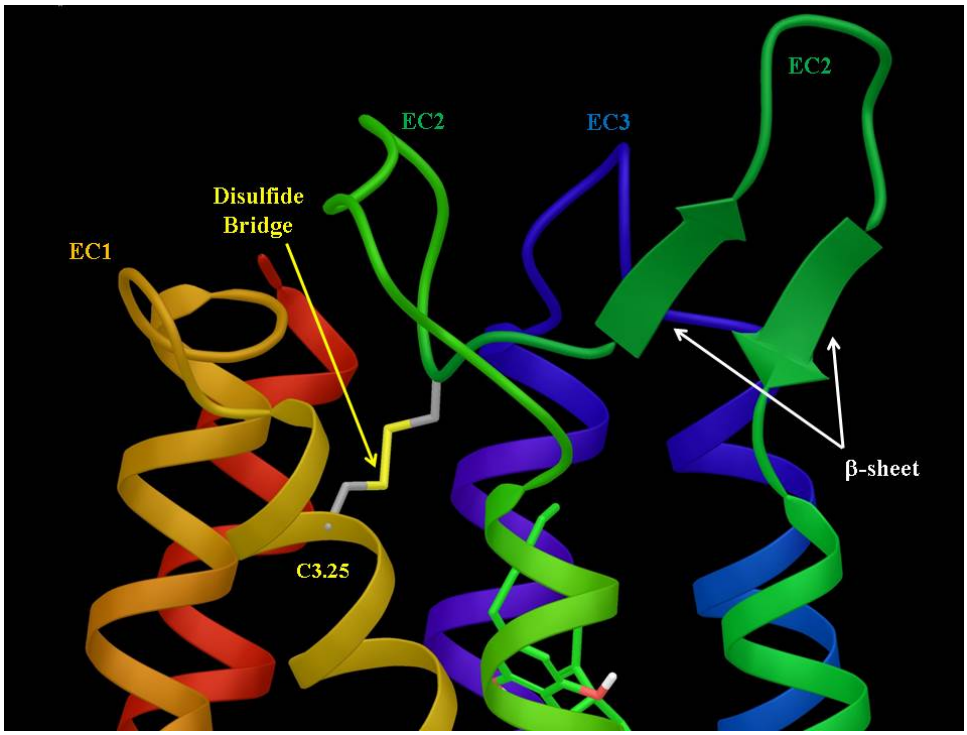


Figure 17. Extracellular Loops of GPR18 Model

Intracellular (IC) Loops

The intracellular loops were modeled and attached in the same way as the extracellular loops, mentioned above, as well as using the same Modeler settings. Unlike the extracellular loops, there were no structural constraints applied to the three loops. GPR18 has a fairly short IC1 loop, containing only five residues. In many GPCRs, the IC2 loop tends to contain a small helical portion (for an example, see the beta-2 adrenergic receptor structure)[9], but due to lack of information, no region was specified

for the calculations. Interestingly, though, Modeler showed a helical extension to TMH3 in a majority of the outputs, despite the proline at 3.57 that usually breaks the helix into loop. This extension also limits the formation of another helix within the middle of the loop, as seen in other crystal structures. It is possible that when placed into a lipid simulation, the random coil of the loops may form into more defined structures.

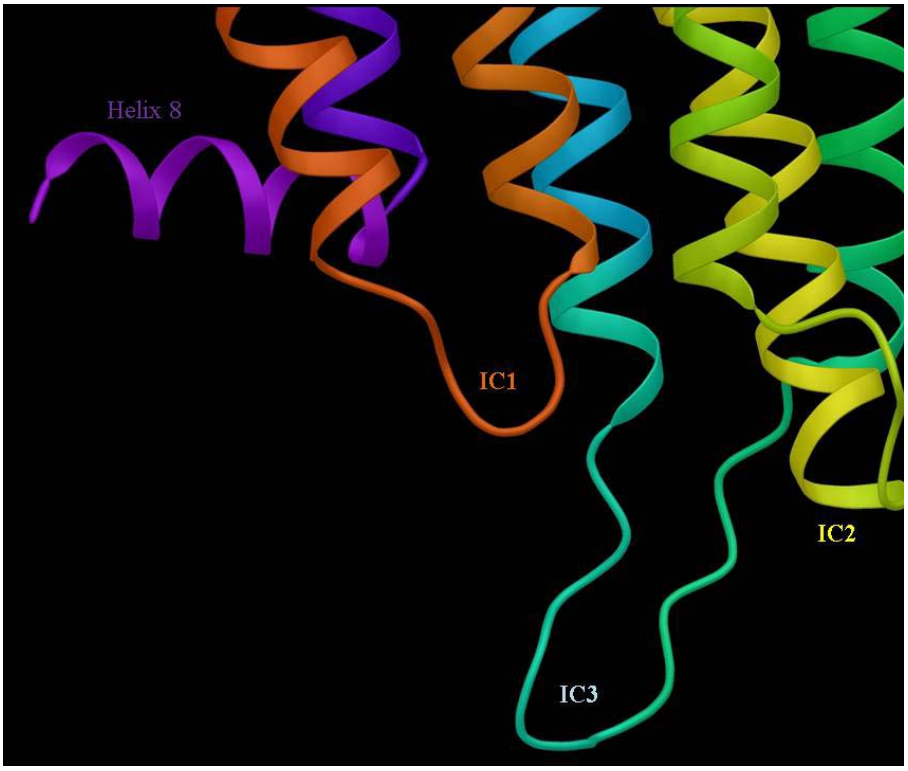


Figure 18. Intracellular Loops of GPR18 Model

In the majority of Class A GPCR crystal structures, the IC3 loop has been broken for the addition of a T4-lysozyme to help stabilize the receptor. Structural information about this

loop is limited, however, no such modification is made in the rhodopsin structure. Here the IC3 loop contains both TMH5 and TMH6 intracellular helical extensions.[7]

Termini Construction and Modeling:

Both the N and C termini were constructed in Maestro and attached to the GPR18 Bundle once both sets of loops were chosen. The N Terminus (18 residues) was run in one calculation, while the C Terminus's length (29 residues) meant that it needed to be run in two smaller sections. Again, a higher dielectric (around 80) was used to simulate the aqueous environment surrounding the loop regions. No structural restrictions were put on the N terminal segment. Using the example of the S1PR1 crystal structure[15] and the proposed method for ligand access to the binding pocket via the lipid bilayer, such as in CB_2,[22] an output was looked for that would cap the top of the bundle. Although no structure was specified, one output revealed a small helical portion that covers the top of the bundle over TMH2 and 3. This section, along with the EC2 loop, provides an almost full cover for the binding pocket.

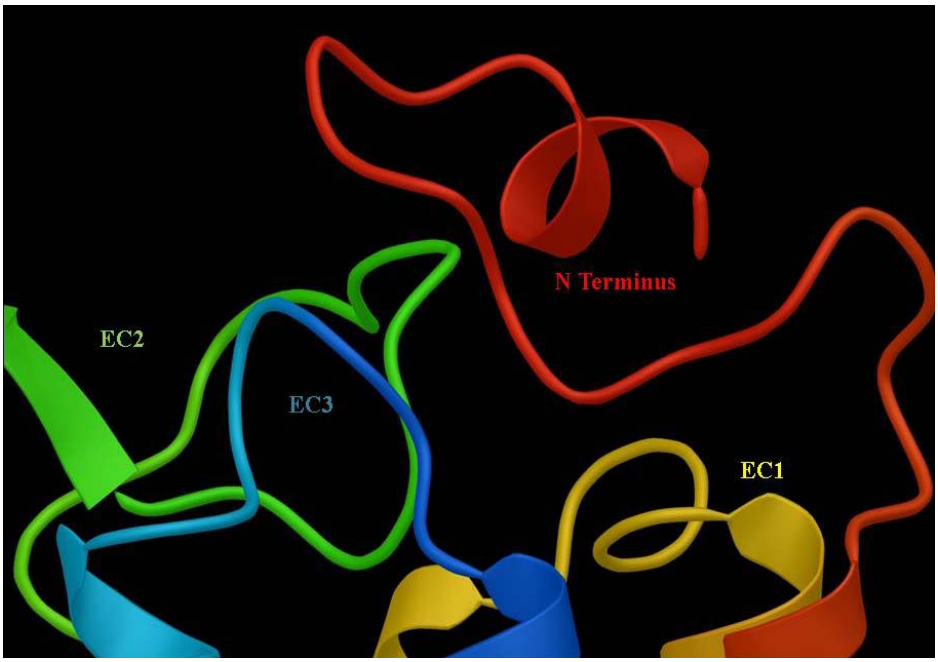


Figure 19. Side view of the N Terminus with its helical portion

The C Terminus (past Helix 8) is complicated as there is little to no structural information about it, from any crystal structure. As there is no cysteine within the C terminus of GPR18, there is no palmitoylation site to be anchored near the bilayer. With a sequence longer than fifteen residues, the terminus was split into two segments, with one run finished and a result chosen before adding the second section. Many of the outputs were culled as the resulting structure either wrapped around the bundle or tried to insert into the bottom of the receptor. The chosen output had the main bulk of the loop situated lower than Helix 8, and the curve in it supported by pi-pi stacking interactions between the multiple arginines and tyrosines within the tail section. Once the first part was chosen, the second was added and run with the same dielectric value.

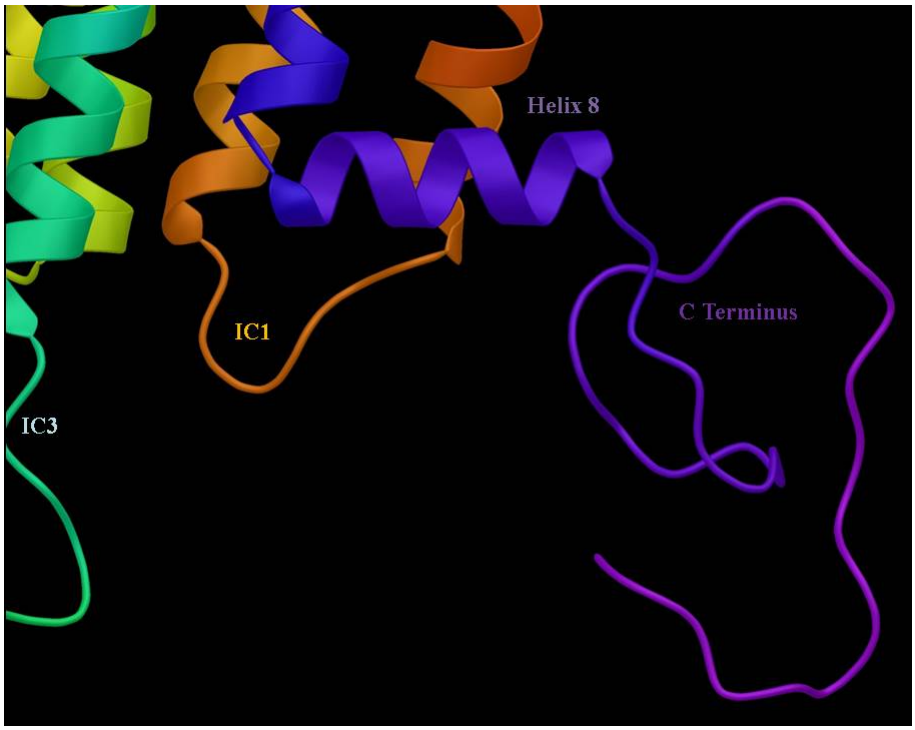


Figure 20. Side view of the C Terminus

The final chosen structure is fairly compact, but not overly tight, so that it would not impede itself when unwinding for later signaling. It also sits largely intracellularly, to help facilitate recognition for signaling.

Interaction Energies:

The interaction energies between CBD and the residues of the binding pocket were calculated using the final minimized dock. The electrostatic, Van der Waals (VDW), and total energy of all residues within 5Å of the ligand were calculated and are shown in Table 1. R2.60 contributes the highest electrostatic interaction of -17.3143

kJ/mol, while R5.42 contributes a smaller, but still significant interaction of -9.6137 kJ/mol. T3.32, F6.51, F6.55, and M7.42 all contribute largely to Van der Waal's interactions, with F6.55 being the highest at around -30 kJ/mol.

Table 1. Interaction Energies between the Ligand CBD and Residues within 5Å

Residue	VDW (kJ/mol)	Electrostatic (kJ/mol)	Total Energy (kJ/mol)
R 2.60	-3.2531	-17.3143	-20.5673
C 3.25	-0.8398	-0.376	-1.2159
L 3.28	-0.7336	0.2813	-0.4523
G 3.29	-6.3752	1.2833	-5.0919
A 3.30	-0.3713	0.0048	-0.3665
T 3.32	-10.0272	-1.257	-11.2842
V 3.33	-8.6347	-0.2797	-8.9145
P 3.36	-2.7599	0.1115	-2.6484
Y 4.64	-4.793	-0.1069	-4.8999
C (EC2)	-0.3639	0.1209	-0.243
R 5.42	-3.6639	-9.6137	-13.2776
F 5.46	-4.0412	0.1351	-3.9062
C 6.47	-0.609	-0.102	-0.711
F 6.48	-6.1463	0.1555	-6.3019
F 6.51	-19.0247	-0.3412	-19.366
H 6.52	-8.1117	-0.1239	-8.2356
C 6.54	-0.7314	0.044	-0.6874
F 6.55	-30.2268	-4.6937	-34.9205
L 6.58	-4.5781	0.1418	-4.4363
M 6.59	-2.3021	-0.0585	-2.3606
T (EC3)	-0.3173	-0.4743	-0.7916
T 7.38	-4.9066	-0.2668	-5.1734
T 7.39	-0.541	-0.0981	-0.6391
M 7.42	-16.6693	-1.5228	-18.1921
			Total Energy (kJ/mol) **-174.6832**

Summary:

A homology model of the inactive state of the orphan GPCR GPR18 was constructed using the recently released crystal structure of the μ-opioid receptor as a base template. Important structural differences in TMH3, 4, and 7 meant reconstructing these helices using Conformational Memories, as well as calculating both active and inactive conformations for TMH6. An ionic lock between S6.33 and the conserved residue of R3.50 was found to stabilize the inactive state of GPR18. Using the lowest energy conformation of the antagonist CBD for docking studies, key interactions with both R2.60 and R5.42 were identified after energy minimization of the CBD/GPR18 R complex. In its final docked conformation, CBD blocks the conformational change of F6.48, thereby acting an antagonist. In agreement with the S1PR1 crystal structure and the hypothesis of a transmembrane entrance into the binding pocket, the modeled N terminus and EC2 loop cover the EC domain of the bundle, thus blocking the ligand from entering the pocket extracellularly.

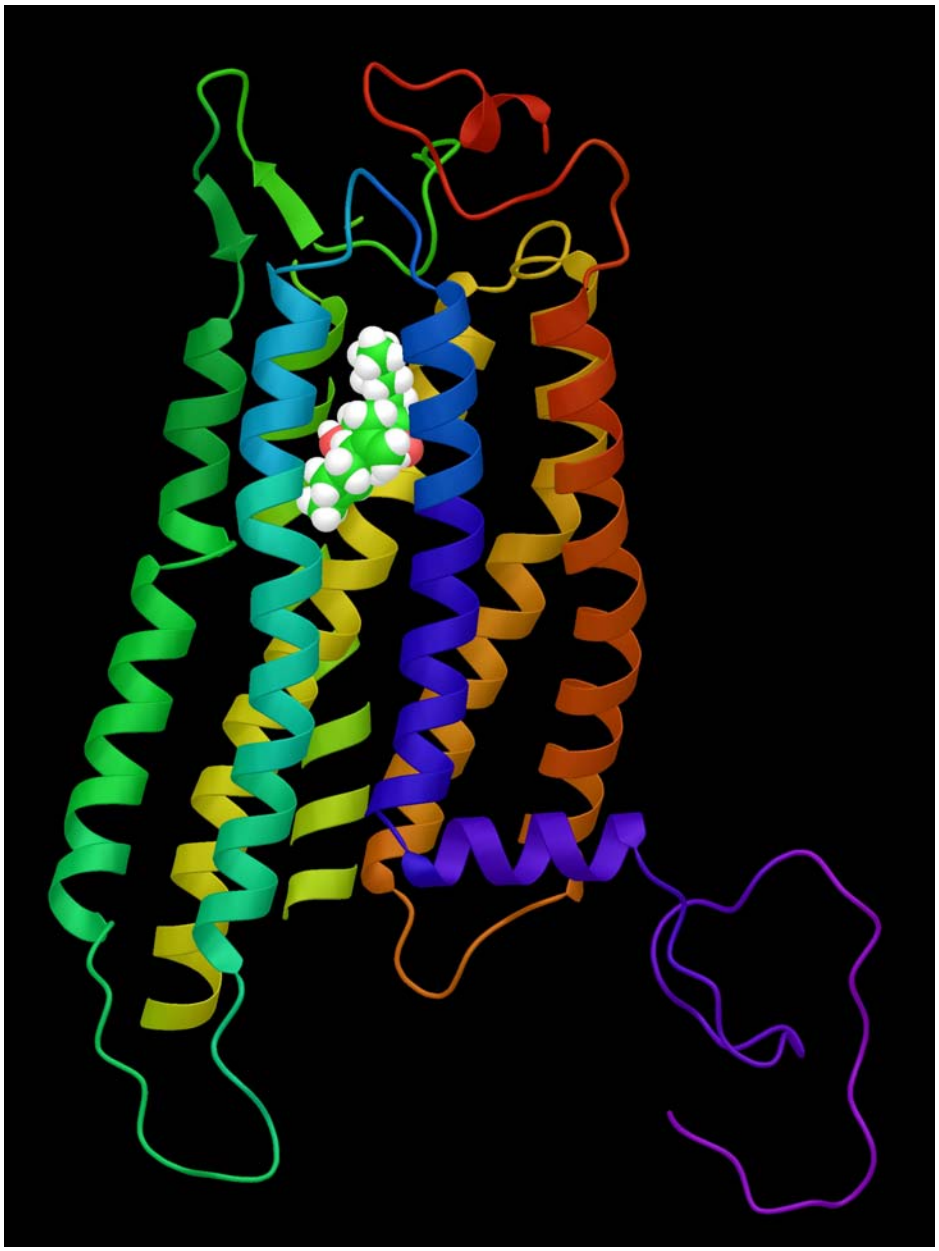


Figure 21. Transmembrane view of complete GPR18 Model

REFERENCES

1. McHugh, D.; Hu, S. S.; Rimmerman, N.; Juknat, A.; Vogel, Z.; Walker, J. M.; Bradshaw, H. B., N-arachidonoyl glycine, an abundant endogenous lipid, potently drives directed cellular migration through GPR18, the putative abnormal cannabidiol receptor. *BMC Neurosci* **2010,** *11*, 44-56.
2. McHugh, D.; Page, J.; Dunn, E.; Bradshaw, H. B., Delta(9) -Tetrahydrocannabinol and N-arachidonyl glycine are full agonists at GPR18 receptors and induce migration in human endometrial HEC-1B cells. *British journal of pharmacology* **2012,** *165* (8), 2414-24.
3. McHugh, D.; Wager-Miller, J.; Page, J.; Bradshaw, H. B., siRNA knockdown of GPR18 receptors in BV-2 microglia attenuates N-arachidonoyl glycine-induced cell migration. *Journal of molecular signaling* **2012,** *7* (1), 10-15.
4. Stella, N.; Sorensen, R.; Abood, M., *Cannbinoids: Actions at Non-CB1/CB2 Cannbinoids Receptors*. Springer Science.
5. Weis, W. I.; Kobilka, B. K., Structural insights into G-protein-coupled receptor activation. *Current Opinion Structural Biology* **2008,** *18* (6), 734-40.
6. Ballesteros, J. A.; Weinstein, H., Integrated Methods for the Construction of Three-Dimensional Probing of Structure-Function Relations in G Protein-Coupled Receptors. In *Methods in Neurosciences*, Sealfon, S. C., Ed. 1995; Vol. 25, pp 366-429.
7. Li, J.; Edwards, P. C.; Burghammer, M.; Villa, C.; Schertler, G. F., Structure of bovine rhodopsin in a trigonal crystal form. *Journal of molecular biology* **2004,** *343* (5), 1409-38.
8. Choe, H. W.; Kim, Y. J.; Park, J. H.; Morizumi, T.; Pai, E. F.; Krauss, N.; Hofmann, K. P.; Scheerer, P.; Ernst, O. P., Crystal structure of metarhodopsin II. *Nature* **2011,** *471* (7340), 651-5.
9. Cherezov, V.; Rosenbaum, D. M.; Hanson, M. A.; Rasmussen, S. G.; Thian, F. S.; Kobilka, T. S.; Choi, H. J.; Kuhn, P.; Weis, W. I.; Kobilka, B. K.; Stevens, R. C., High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science* **2007,** *318* (5854), 1258-65.
10. Warne, T.; Serrano-Vega, M. J.; Baker, J. G.; Moukhametzianov, R.; Edwards, P. C.; Henderson, R.; Leslie, A. G.; Tate, C. G.; Schertler, G. F., Structure of a beta1-adrenergic G-protein-coupled receptor. *Nature* **2008,** *454* (7203), 486-91.
11. Jaakola, V. P.; Griffith, M. T.; Hanson, M. A.; Cherezov, V.; Chien, E. Y.; Lane, J. R.; Ijzerman, A. P.; Stevens, R. C., The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist. *Science* **2008,** *322* (5905), 1211-7.
12. Wu, B.; Chien, E. Y.; Mol, C. D.; Fenalti, G.; Liu, W.; Katritch, V.; Abagyan, R.; Brooun, A.; Wells, P.; Bi, F. C.; Hamel, D. J.; Kuhn, P.; Handel, T. M.; Cherezov, V.; Stevens, R. C., Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science* **2010,** *330* (6007), 1066-71.

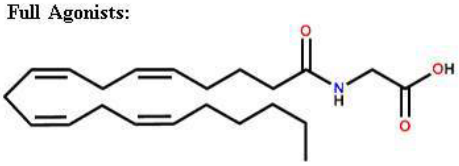
13. Chien, E. Y.; Liu, W.; Zhao, Q.; Katritch, V.; Han, G. W.; Hanson, M. A.; Shi, L.; Newman, A. H.; Javitch, J. A.; Cherezov, V.; Stevens, R. C., Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. *Science* **2010,** *330* (6007), 1091-5.
14. Shimamura, T.; Shiroishi, M.; Weyand, S.; Tsujimoto, H.; Winter, G.; Katritch, V.; Abagyan, R.; Cherezov, V.; Liu, W.; Han, G. W.; Kobayashi, T.; Stevens, R. C.; Iwata, S., Structure of the human histamine H1 receptor complex with doxepin. *Nature* **2011,** *475* (7354), 65-70.
15. Hanson, M. A.; Roth, C. B.; Jo, E.; Griffith, M. T.; Scott, F. L.; Reinhart, G.; Desale, H.; Clemons, B.; Cahalan, S. M.; Schuerer, S. C.; Sanna, M. G.; Han, G. W.; Kuhn, P.; Rosen, H.; Stevens, R. C., Crystal structure of a lipid G protein-coupled receptor. *Science* **2012,** *335* (6070), 851-5.
16. Thompson, A. A.; Liu, W.; Chun, E.; Katritch, V.; Wu, H.; Vardy, E.; Huang, X. P.; Trapella, C.; Guerrini, R.; Calo, G.; Roth, B. L.; Cherezov, V.; Stevens, R. C., Structure of the nociceptin/orphanin FQ receptor in complex with a peptide mimetic. *Nature* **2012,** *485* (7398), 395-9.
17. Manglik, A.; Kruse, A. C.; Kobilka, T. S.; Thian, F. S.; Mathiesen, J. M.; Sunahara, R. K.; Pardo, L.; Weis, W. I.; Kobilka, B. K.; Granier, S., Crystal structure of the μ-opioid receptor bound to a morphinan antagonist. *Nature* **2012,** *485* (7398), 321-6.
18. Granier, S.; Manglik, A.; Kruse, A. C.; Kobilka, T. S.; Thian, F. S.; Weis, W. I.; Kobilka, B. K., Structure of the delta-opioid receptor bound to naltrindole. *Nature* **2012,** *485* (7398), 400-4.
19. Wu, H.; Wacker, D.; Mileni, M.; Katritch, V.; Han, G. W.; Vardy, E.; Liu, W.; Thompson, A. A.; Huang, X. P.; Carroll, F. I.; Mascarella, S. W.; Westkaemper, R. B.; Mosier, P. D.; Roth, B. L.; Cherezov, V.; Stevens, R. C., Structure of the human kappa-opioid receptor in complex with JDTic. *Nature* **2012,** *485* (7398), 327-32.
20. Kohno, M.; Hasegawa, H.; Inoue, A.; Muraoka, M.; Miyazaki, T.; Oka, K.; Yasukawa, M., Identification of N-arachidonylglycine as the endogenous ligand for orphan G-protein-coupled receptor GPR18. *Biochemical and Biophysical Research Communications* **2006,** *347* (3), 827-32.
21. Pei, Y.; Mercier, R. W.; Anday, J. K.; Thakur, G. A.; Zvonok, A. M.; Hurst, D.; Reggio, P. H.; Janero, D. R.; Makriyannis, A., Ligand-binding architecture of human CB2 cannabinoid receptor: evidence for receptor subtype-specific binding motif and modeling GPCR activation. *Chemistry & biology* **2008,** *15* (11), 1207-19.
22. Hurst, D. P.; Grossfield, A.; Lynch, D. L.; Feller, S.; Romo, T. D.; Gawrisch, K.; Pitman, M. C.; Reggio, P. H., A lipid pathway for ligand binding is necessary for a cannabinoid G protein-coupled receptor. *The Journal of biological chemistry* **2010,** *285* (23), 17954-64.
23. Farrens, D. L.; Altenbach, C.; Yang, K.; Hubbell, W. L.; Khorana, H. G., Requirement of rigid-body motion of transmembrane helices for light activation of rhodopsin. *Science* **1996,** *274* (5288), 768-70.
24. Shi, L.; Liapakis, G.; Xu, R.; Guarnieri, F.; Ballesteros, J. A.; Javitch, J. A., Beta2

adrenergic receptor activation. Modulation of the proline kink in transmembrane 6 by a rotamer toggle switch. *The Journal of biological chemistry* **2002,** *277* (43), 40989-96.
25. Ghanouni, P.; Steenhuis, J. J.; Farrens, D. L.; Kobilka, B. K., Agonist-induced conformational changes in the G-protein-coupling domain of the beta 2 adrenergic receptor. *Proceedings of the National Academy of Sciences of the United States of America* **2001,** *98* (11), 5997-6002.
26. Hamm, H. E.; Deretic, D.; Arendt, A.; Hargrave, P. A.; Koenig, B.; Hofmann, K. P., Site of G protein binding to rhodopsin mapped with synthetic peptides from the alpha subunit. *Science* **1988,** *241* (4867), 832-5.
27. Huang, P.; Visiers, I.; Weinstein, H.; Liu-Chen, L. Y., The local environment at the cytoplasmic end of TM6 of the mu opioid receptor differs from those of rhodopsin and monoamine receptors: introduction of an ionic lock between the cytoplasmic ends of helices 3 and 6 by a L6.30(275)E mutation inactivates the mu opioid receptor and reduces the constitutive activity of its T6.34(279)K mutant. *Biochemistry* **2002,** *41* (40), 11972-80.
28. Whitnell, R. M.; Hurst, D. P.; Reggio, P. H.; Guarnieri, F., Conformational memories with variable bond angles. *Journal of Computational Chemistry* **2008,** *29* (5), 741-52.
29. Zhang, R.; Hurst, D. P.; Barnett-Norris, J.; Reggio, P. H.; Song, Z. H., Cysteine 2.59(89) in the second transmembrane domain of human CB2 receptor is accessible within the ligand binding crevice: evidence for possible CB2 deviation from a rhodopsin template. *Molecular pharmacology* **2005,** *68* (1), 69-83.
30. Visiers, I.; Braunheim, B. B.; Weinstein, H., Prokink: a protocol for numerical evaluation of helix distortions by proline. *Protein engineering* **2000,** *13* (9), 603-6.
31. Kotsikorou, E.; Madrigal, K. E.; Hurst, D. P.; Sharir, H.; Lynch, D. L.; Heynen-Genel, S.; Milan, L. B.; Chung, T. D.; Seltzman, H. H.; Bai, Y.; Caron, M. G.; Barak, L.; Abood, M. E.; Reggio, P. H., Identification of the GPR55 agonist binding site using a novel set of high-potency GPR55 selective ligands. *Biochemistry* **2011,** *50* (25), 5633-47.
32. Fiser, A.; Do, R. K.; Sali, A., Modeling of loops in protein structures. *Protein science : a publication of the Protein Society* **2000,** *9* (9), 1753-73.
33. Console-Bram, L. B., Eugen; Brailoiu, Cristina; Abood, Mary E. In *Calcium Mobilization and MAPK Activity via GPR18*, 57th Annual Meeting Biophysical Society, Philadelphia, Pennsylvania, February 5, 2013; Philadelphia, Pennsylvania, 2013.

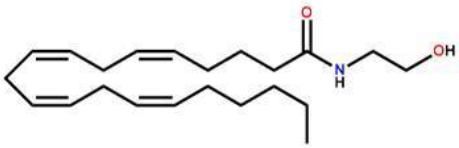
APPENDIX A

GPR18 LIGANDS

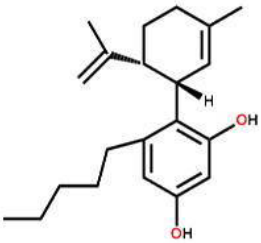
Full Agonists:



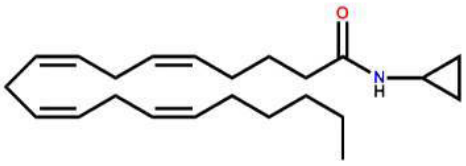
NAGly



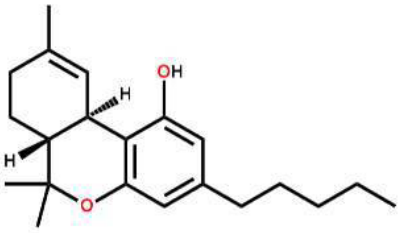
AEA



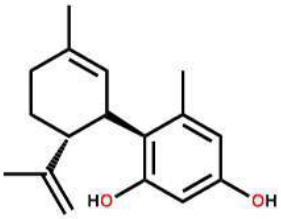
Abn-CBD



ACPA

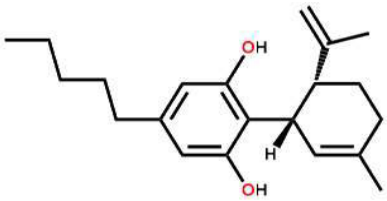


Δ^9-THC

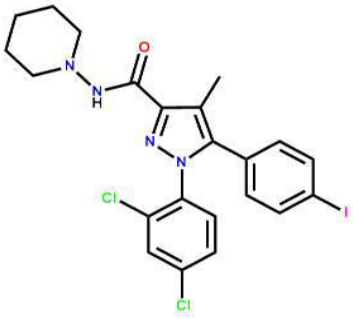


O-1602

Partial Agonists/Antagonists:



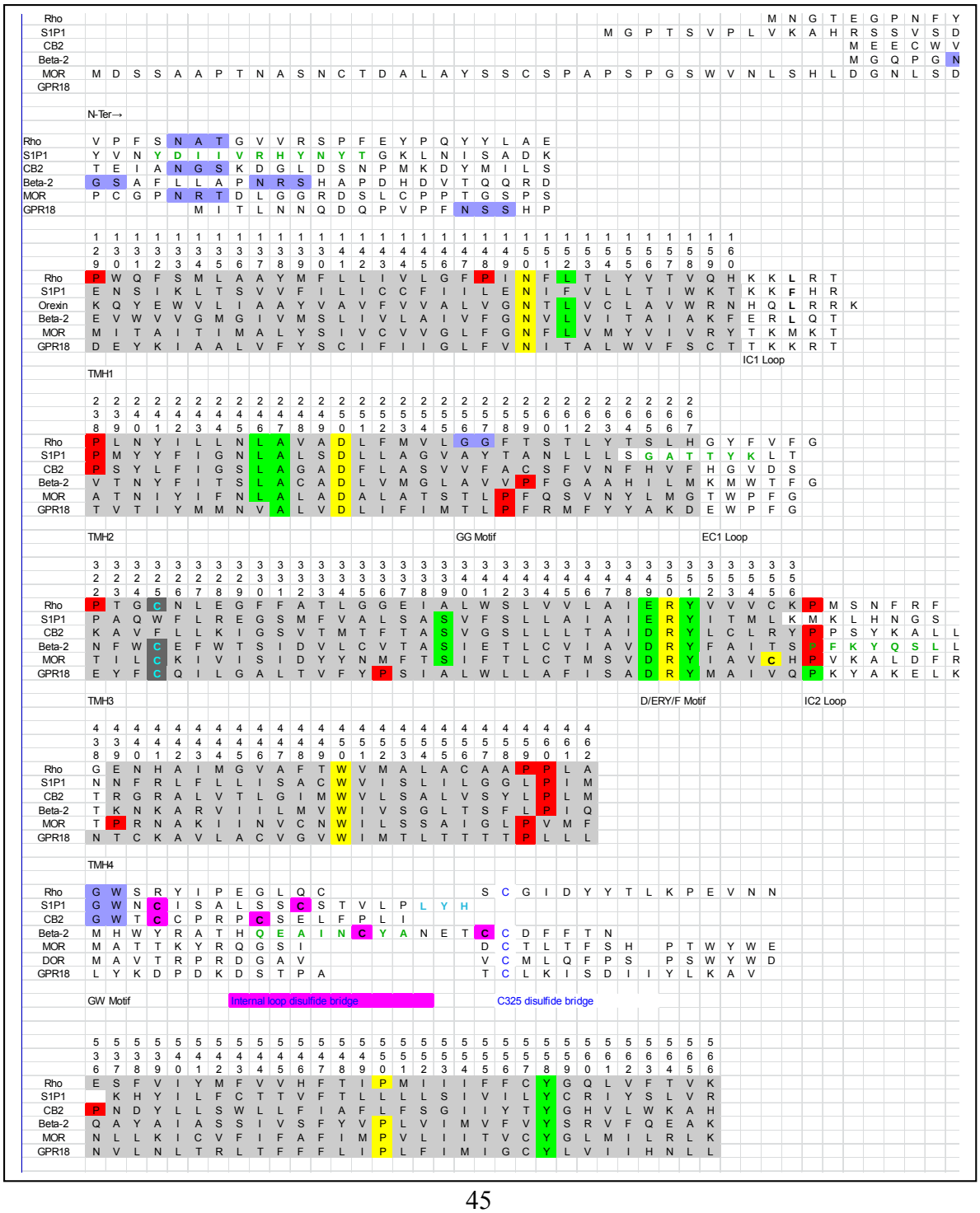
CBD

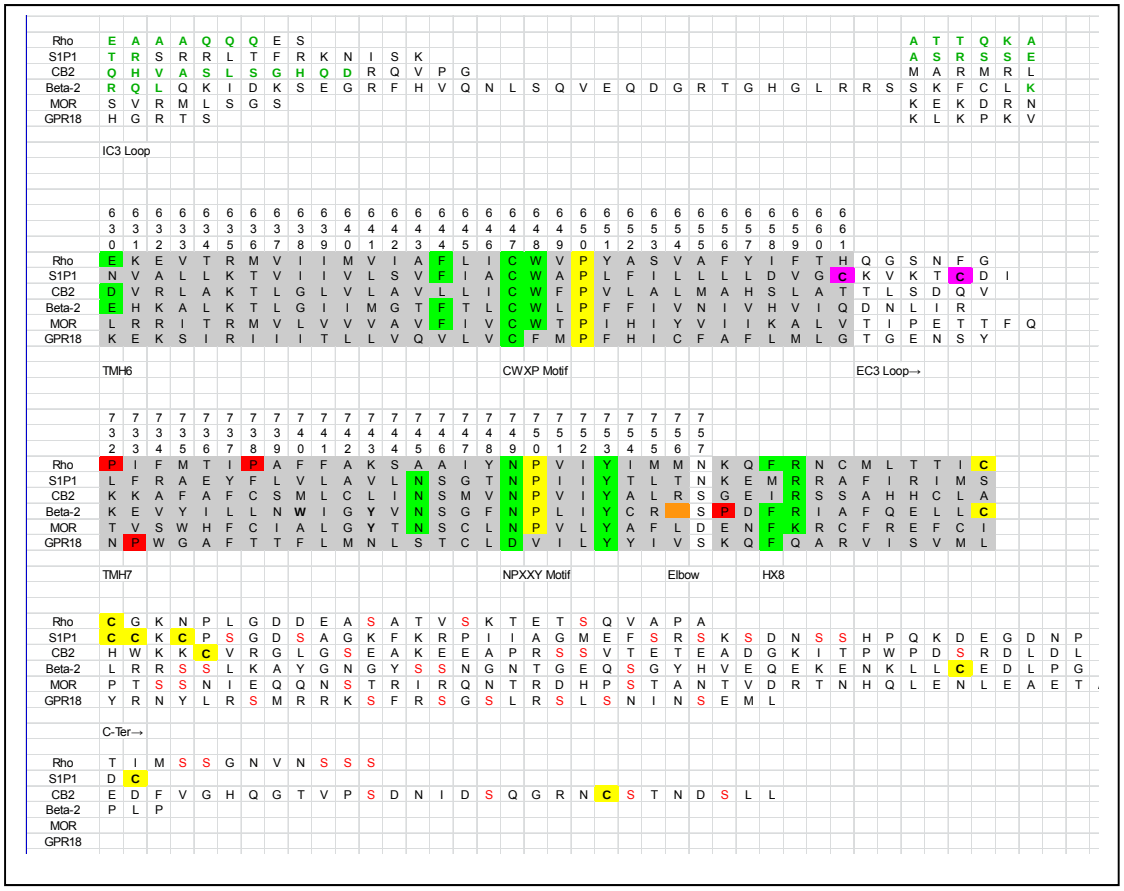


AM251

APPENDIX B

SEQUENCE ALIGNMENT




IC3 Loop
TMH6
CWXP Motif
EC3 Loop→
TMH7
NPXXY Motif
Elbow
HX8
C-Ter→